



Optimal Camera Placement to Obtain Accurate 3D Point Positions

Gustavo Olague, Roger Mohr

► To cite this version:

Gustavo Olague, Roger Mohr. Optimal Camera Placement to Obtain Accurate 3D Point Positions. International Conference on Pattern Recognition, Aug 1998, Brisbane, Australia. pp.8–10, 10.1109/ICPR.1998.711066 . inria-00548331

HAL Id: inria-00548331

<https://inria.hal.science/inria-00548331>

Submitted on 22 Dec 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optimal Camera Placement to Obtain Accurate 3D Point Positions

Gustavo Olague, Roger Mohr
MOVI-GRAVIR*, ZIRST - 655 Avenue de l'Europe
38330 Montbonnot Saint Martin, France
Gustavo.Olague@inrialpes.fr Roger.Mohr@inrialpes.fr

1. Introduction and Approach

The problem we would like to approach is the automation of the camera network design process in order to obtain accurate 3D measurements. We restrict ourselves to the problem where the camera positions are limited only by the incidence angle constraint and it is simplified to the case where the cameras remain at a fixed distance to the set of target points to be measured. The main question we would like to answer is: Where should we place the cameras in order to obtain the minimal 3D error? From this question several subproblems arise; How can we develop a good criterion to judge our configuration? What conditions are needed for our system to work? Which are the inter-related aspects involved in the development of the system? What would be a good method to optimize the placement of the camera? From these initial questions the choice of a **criterion** combined with an **optimization** process will be the key concepts.

Researchers in the computer vision and photogrammetric communities have recognized the need to automate the process of camera network design. Tarbox and Gottschlich [7] have recognised the need for *multistation solutions* to overcome object occlusion problems. They have implemented a solution in the IVIS system for an active triangulation sensor. Fritsch and Crotilla [3] have investigated the potential of optimizing multi-station configurations using an analytical first order design (FOD) approach by iteratively shifting the sensor stations until the covariance matrix of the estimated object feature coordinates is better than a criterion matrix. However their approach is limited to the optimization of existing configurations. A work called CONSENS which follows the expert system approach and uses multiple cameras in combination with optical triangulation, was developed by Mason et al. [5]. It describes an ideal configuration of four camera stations that can be employed to provide a strong imaging geometry for the class of planar network design problems. Complex objects are di-

vided into planes, each one is evaluated through one of these networks and then connected with some extra cameras with the purpose to establish just one common datum. Unlike all other approaches, our idea is to pose the problem in terms of a global optimization design which is capable of managing the problem using an adaptive strategy. It explores the solution space using both non continuous optimization and combinatorial search.

The approach can be divided into two main components. Firstly, we will develop an analytic uncertainty analysis based on error propagation phenomenon. This will allow us to express an error criterion to be minimized. Secondly, we will present an evolutionary optimization method similar to genetic algorithm [4], which optimizes this criterion.

1.1. Uncertainty Analysis

A camera can be considered as a system that performs a linear projective transformation from the projective space \mathbb{P}^3 into the projective plane \mathbb{P}^2 .

$$p_{ij} \sim M_i P_j . \quad (1)$$

This mapping is represented by a matrix $M_{3 \times 4}$

$$M = K(R \mid T) , \quad (2)$$

where K is the matrix of intrinsic parameters, R is a rotation matrix and T a translation vector. Equation (1) can be rewritten as

$$A(p_i, M_i)P = b(p_i, M_i) , \quad (3)$$

where A is a $2i \times 3$ matrix and b is a $2i \times 1$ vector. If matrix $A^t A$ is invertible we can find the least squares solution

$$P = (A^t A)^{-1} A^t b , \quad (4)$$

which minimizes $\|AP - b\|^2$.

The key of manipulating geometric uncertainty is to be able to transform the information or probability density function on a feature available in one form (image point)

*MOVI is a joint research project between the CNRS, INPG, INRIA Rhône-Alpes and UJF.

into another form of interest (point in space), given by Equation(4), see Faugeras [1]. This transformation of information can be grouped into a family of transformations which we approximate to the exact transformation by a first-order relation using Taylor series. In this way, a linear approximation is to be used in which we assume a Gaussian distribution. Then the mean $E[P]$ and covariance ΛP are sufficient information to completely define the feature density function. All this is given by the following proposition (see [1, chapter V]):

Proposition 1 *Given a random variable $p \in \mathbb{R}^m$, of Gaussian distribution, mean $E[p]$, and covariance Λp , and $P \in \mathbb{R}^n$, the random vector given by $P = f(p)$, where f is a function of class C^1 , the mean of P can be approximated to a first-order Taylor expansion by $f(E[p])$ and its covariance by:*

$$\Lambda P = \frac{\partial f(E[p])}{\partial p} \Lambda p \frac{\partial f(E[p])^t}{\partial p} . \quad (5)$$

Therefore, ΛP is a symmetric positive definite matrix, which describes the bounds on $P = f(p)$ in the vicinity of $E[P] = f(E[p])$ given those of p in the vicinity of $E[p]$.

1.2. The Criteria

Once we have computed the covariance matrix ΛP , it is necessary to choose a criterion useful to the optimization process. In this sense, we need to select a metric to compare symmetric positive definite matrices. The *comparison* of covariance matrices is interpreted as the required standard deviation of function $P = f(p)$ to be better when it is computed with covariance matrix ΛP_1 than with ΛP_2 . This comparison can be achieved using: the maximum eigenvalue of ΛP , the trace of ΛP or a matrix criterion. However, we propose to use the *maximum element in the diagonal* of ΛP which corresponds to the worst variance of individual parameters

$$\mu(P) = \max_{j=1 \dots 3} \Lambda P_{jj} . \quad (6)$$

This measure is easy to compute and gives a good uniformity of network precision in each coordinate axis of the error ellipsoid represented by ΛP .

1.3. The Optimization Process

The problem of camera network design presents discontinuities aspects mainly due to the unobservability of points which leads to a combinatorial optimization process which we have approached using a multicellular genetic algorithm (MGA) [6]. This algorithm proceeds as follows:

1. an initial random population of N convergent networks represented by (α_n, β_n) .
 2. next, we evaluate each network, and store the corresponding maximum value of the diagonal of ΛP_n for each tree structure.
 3. then, we select a population of “good” networks by *tournament selection*.
 4. from this population, we recombine the variables (α_n, β_n) for each camera using the following operations: Crossover, with a probability $P_c = 0.7$; Mutation, with a probability $P_m = 0.005$.
- These operations yields a new population which we copy into $P(t)$.
5. Steps 2,3, and 4 are repeated until the optimization criterion stabilizes.

Finally, this algorithm minimizes the maximum value in the diagonal of ΛP

$$\min_{i=1 \dots N} (\mu(P)) . \quad (7)$$

Thereby, the camera placement M_i relative to the world coordinate frame is optimized. Geometrically, each ΛP_i represents a hyper-ellipsoid, which changes its orientation and size as each sensor placement M_i does. Thus, an optimal placement solution is proposed, where the combined uncertainty of all points is minimal.

2. Experiments

We have run a series of experiments to test the validity of our approach. We present some results in Figure 1 which shows some cases of configurations designed by EPOCA in which several cameras are looking at several targets represented by their error ellipsoids aligned in one, two, or three planes. These configurations are product of our evolutionary system.

As an example the graphs of Figure 1 were produced with a population of 30 camera-networks which converged in no more than 150 generations. These graphs show the maximal, average and best fitness for one run of the algorithm. This experiment took a total of 72 generations to converge, discovering the best configuration with a fitness value = .0047 in generation 59. It has required about 7 seconds of CPU time on an UltraSPARC 200 Mhz. Note also that evolution is terminated near generation 60. Notice the behaviour of the angle α of camera 1 as illustrated in Figures 1e. We see that after a random distribution, emergent coherent configurations appear, then this variable converges towards two dominating values below and above 40 degrees. If we analyse how $\mu(P)$ evolves with these parameters, we will observe the local minima of the dominating

configurations. Figure 1c displays the graph of $\mu(P)$, after convergence, only with respect of two parameters and for camera 1. It clearly shows the two minima for the β parameter of camera 1, see Figure 1d, while the landscape is quite flat for the α parameter. Similar observations can be derived for all the other parameters.

3. Discussion and Perspectives

The problem presents discontinuities which leads to combinatorial aspects in the optimization process. These constraints are naturally incorporated to the genetic algorithm methodology. Our EPOCA system successfully produces two and three camera networks designs similar to those used by photogrammetrists. In the case of four cameras, see Figure 1f, a non-standard design was proposed. Photogrammetrists usually put the four camera at four corners of a cube whose center contains the targets to be measured. In fact our configuration was already pointed out by Fraser [2]; he noticed that this configuration is not atypical. Our experiments confirm Fraser statement. Moreover, the system can design networks for several adjoining planes.

Acknowledgments Gustavo Olague is grateful for research founding of CONACYT México (grant 66291). Figure 1f was generated with software written at the Geometry Center, University of Minnesota.

References

- [1] O. Faugeras. *Three-Dimensional Computer Vision - A Geometric Viewpoint*. Artificial intelligence. The MIT Press, Cambridge, MA, USA, Cambridge, MA, 1993.
- [2] C. Fraser. Optimization of precision in close-range photogrammetry. *Photogrammetric Engineering and Remote Sensing*, 48(4):561–570, April 1982.
- [3] D. Fritsch and F. Crosilla. First order design strategies for industrial photogrammetry. In *Close-Range Photogrammetry Meets Machine Vision*, volume 1395, pages 432–438. SPIE - Society of Photo-Optical Instrumentation Engineers, 1990.
- [4] J. Holland. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*. The MIT Press, Cambridge, MA, USA, 1992.
- [5] S. Mason and A. Gruen. Automatic sensor placement for accurate dimensional inspection. *Computer Vision and Image Understanding*, 3(61):454–467, 1995.
- [6] G. Olague and R. Mohr. Optimal camera placement for accurate reconstruction. Technical Report 3338, INRIA Rhône Alpes, January 1998. ftp://ftp.imag.fr/pub/MOVI/technical_reports/RR-3338.ps.gz.
- [7] G. Tarbox and S. Gottschlich. Planning for complete sensor coverage in inspection. Technical Report TR-CAT-93-4, NYS Center for Advanced Technology in Automation and Robotics., 1993.

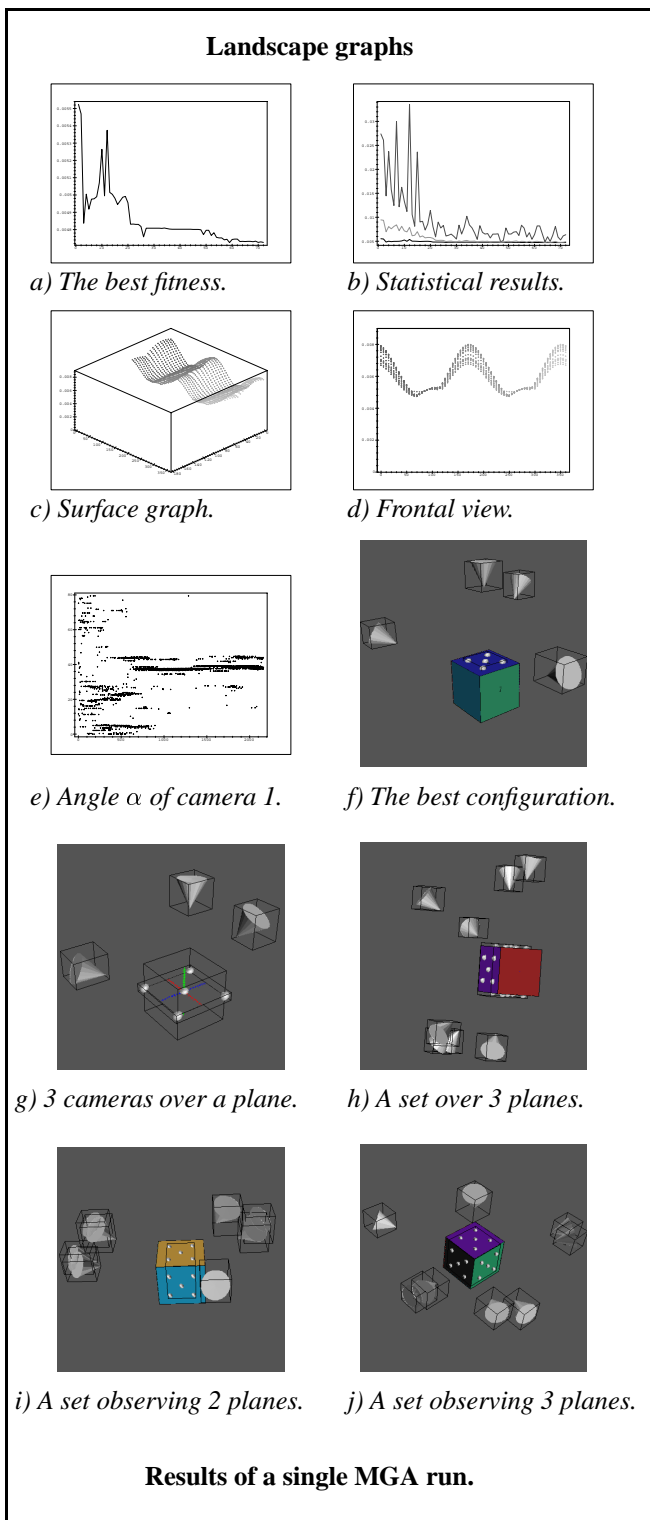


Figure 1. The following graphs show the statistics and convergence time of a single MGA run composed of 4 convergent cameras.